

## DAFTAR ISI

|  |      |
|--|------|
| PERNYATAAN .....                           | ii   |
| PRAKATA .....                              | iii  |
| ARTI LAMBANG DAN SINGKATAN .....           | v    |
| ABSTRACT .....                             | vi   |
| INTISARI.....                              | vii  |
| DAFTAR ISI .....                           | viii |
| DAFTAR GAMBAR .....                        | x    |
| DAFTAR TABEL.....                          | xi   |
| BABI PENDAHULUAN                           |      |
| 1.1 Latar Belakang.....                    | 1    |
| 1.2 Perumusan masalah.....                 | 2    |
| 1.3 Keaslian penelitian .....              | 2    |
| 1.4 Tujuan Penelitian.....                 | 8    |
| 1.5 Manfaat Penelitian .....               | 8    |
| BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI |      |
| 2.1 Klasifikasi Halaman Web .....          | 9    |
| 2.2 Proses <i>Preprocessing</i> .....      | 10   |
| 2.3 Proses Seleksi Fitur .....             | 11   |
| 2.4 Algoritme Pengklasifikasi .....        | 22   |
| 2.5 <i>K-Fold Cross Validation</i> .....   | 28   |
| 2.6 Pertanyaan Penelitian .....            | 29   |
| BAB III METODOLOGI                         |      |
| 3.1 Alat dan Bahan.....                    | 30   |
| 3.2 Alur Penelitian .....                  | 32   |
| 3.3 Perancangan Konseptual Sistem .....    | 34   |
| 3.4 Metode Analisis.....                   | 39   |
| 3.5 Skenario Pengujian .....               | 40   |

## BAB IV HASIL DAN PEMBAHASAN

4.1 Hasil Pengujian Seleksi Fitur ..... 41

4.2 Hasil Pengujian Klasifikasi ..... 42

## BAB V KESIMPULAN DAN SARAN

5.1 Kesimpulan ..... 53

5.2 Saran ..... 54

DAFTAR PUSTAKA ..... 55

## LAMPIRAN

Lampiran 1 ..... L-58

Lampiran 2 ..... L-59

Lampiran 3 ..... L-60

## DAFTAR GAMBAR

|            |  |    |
|------------|--|----|
| Gambar 2.1 | Proses Klasifikasi .....                       | 9  |
| Gambar 2.2 | Proses <i>Tokenizing</i> .....                 | 11 |
| Gambar 2.3 | <i>Stoplist</i> .....                          | 11 |
| Gambar 2.4 | Analogi Proses Seleksi Fitur.....              | 12 |
| Gambar 2.5 | Siklus Algoritme Genetika .....                | 14 |
| Gambar 2.6 | Roda Lotere .....                              | 16 |
| Gambar 2.7 | <i>Support Vector Machine</i> .....            | 17 |
| Gambar 2.8 | Siklus Algoritme Decision Trees .....          | 24 |
| Gambar 2.9 | Proses Fold Cross Validation .....             | 28 |
| Gambar 3.1 | Contoh Dataset WebKb .....                     | 31 |
| Gambar 3.2 | Contoh Dataset Router (R8) .....               | 32 |
| Gambar 3.3 | Diagram Alur Penelitian .....                  | 32 |
| Gambar 3.4 | Alur Proses Sistem .....                       | 34 |
| Gambar 3.5 | Seleksi Fitur Menggunakan GA .....             | 36 |
| Gambar 3.6 | Seleksi Fitur Menggunakan <i>Wrapper</i> ..... | 37 |
| Gambar 3.7 | Proses Klasifikasi .....                       | 38 |
| Gambar 3.8 | Komparasi Proses Klasifikasi .....             | 39 |
| Gambar 4.1 | Hasil Seleksi Fitur .....                      | 41 |

## DAFTAR TABEL

|           |   |    |
|-----------|---|----|
| Tabel 1.1 | Rangkuman Keaslian Penelitian .....               | 6  |
| Tabel 1.2 | Rangkuman Keaslian Penelitian (Lanjutan) .....    | 7  |
| Tabel 2.1 | Contoh Perhitungan Algoritme J48 .....            | 25 |
| Tabel 2.2 | Contoh Perhitungan Algoritme J48 (Lanjutan) ..... | 26 |
| Tabel 2.3 | Tabel Atribut Data .....                          | 26 |
| Tabel 2.4 | Hasil Rata-Rata <i>Entropy</i> .....              | 28 |
| Tabel 3.1 | Dataset Penelitian .....                          | 31 |
| Tabel 4.1 | Nilai Presisi WebKb .....                         | 43 |
| Tabel 4.2 | Nilai Presisi Router (r8) .....                   | 45 |
| Tabel 4.3 | Nilai Recall WebKb .....                          | 48 |
| Tabel 4.4 | Nilai Recall Router (r8) .....                    | 49 |
| Tabel 4.5 | Nilai Rata-Rata Akurasi .....                     | 51 |